

The 3rd Annual Health Econometrics Workshop



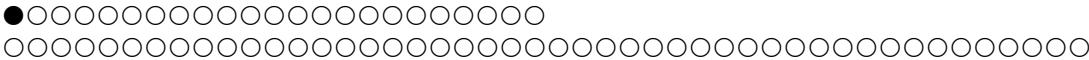
Latent Class Modeling

William Greene
Department of Economics
Stern School of Business
New York University

3rd Annual Health Econometrics Workshop
University of Minnesota, Minneapolis, MN
September 29-October 1, 2011

Outline

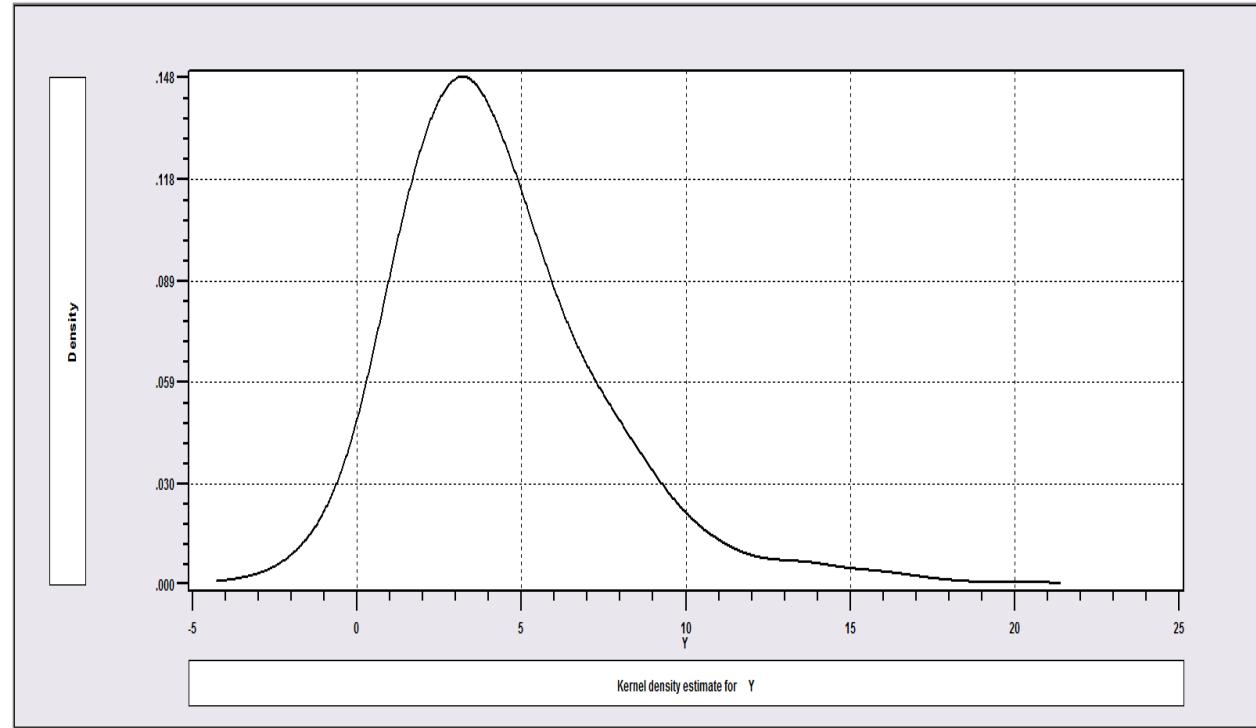
- Finite mixture and latent class models
- Extensions of the latent class model
- Applications of several variations
 - Health economics
 - Transport
 - Production and efficiency
- Aspects of estimation of LC models



The “Finite Mixture Model”

- An unknown parametric model governs an outcome y
 - $F(y|\mathbf{x}, \theta)$
 - This is *the* model
 - We approximate $F(y|\mathbf{x}, \theta)$ with a weighted sum of specified (e.g., normal) densities:
 - $$F(y|\mathbf{x}, \theta) \approx \sum_j \pi_j G(y|\mathbf{x}, \beta)$$
 - This is a search for functional form. With a sufficient number of (normal) components, we can approximate any density to any desired degree of accuracy. (McLachlan and Peel (2000))
 - There is no “mixing” ***process*** at work

	Y
1 »	4.21855
2 »	1.20367
3 »	2.45719
4 »	0.470427
5 »	16.4708
6 »	0.428376
7 »	1.56961
8 »	5.93268
9 »	3.83085
10 »	4.10209
11 »	7.29334
12 »	14.278
13 »	9.12016
14 »	1.57473
15 »	5.19982
16 »	3.84372
17 »	-3.57989
18 »	2.32862
19 »	2.85411
20 »	5.23678
21 »	2.25915
22 »	3.22748
23 »	11.0248
24 »	4.31525
25 »	3.55592
26 »	7.30238
27 »	8.61563
28 »	1.31486
29 »	5.6779
30 »	11.3807



Density? Note significant mass below zero. Not a gamma or lognormal or any other familiar density.



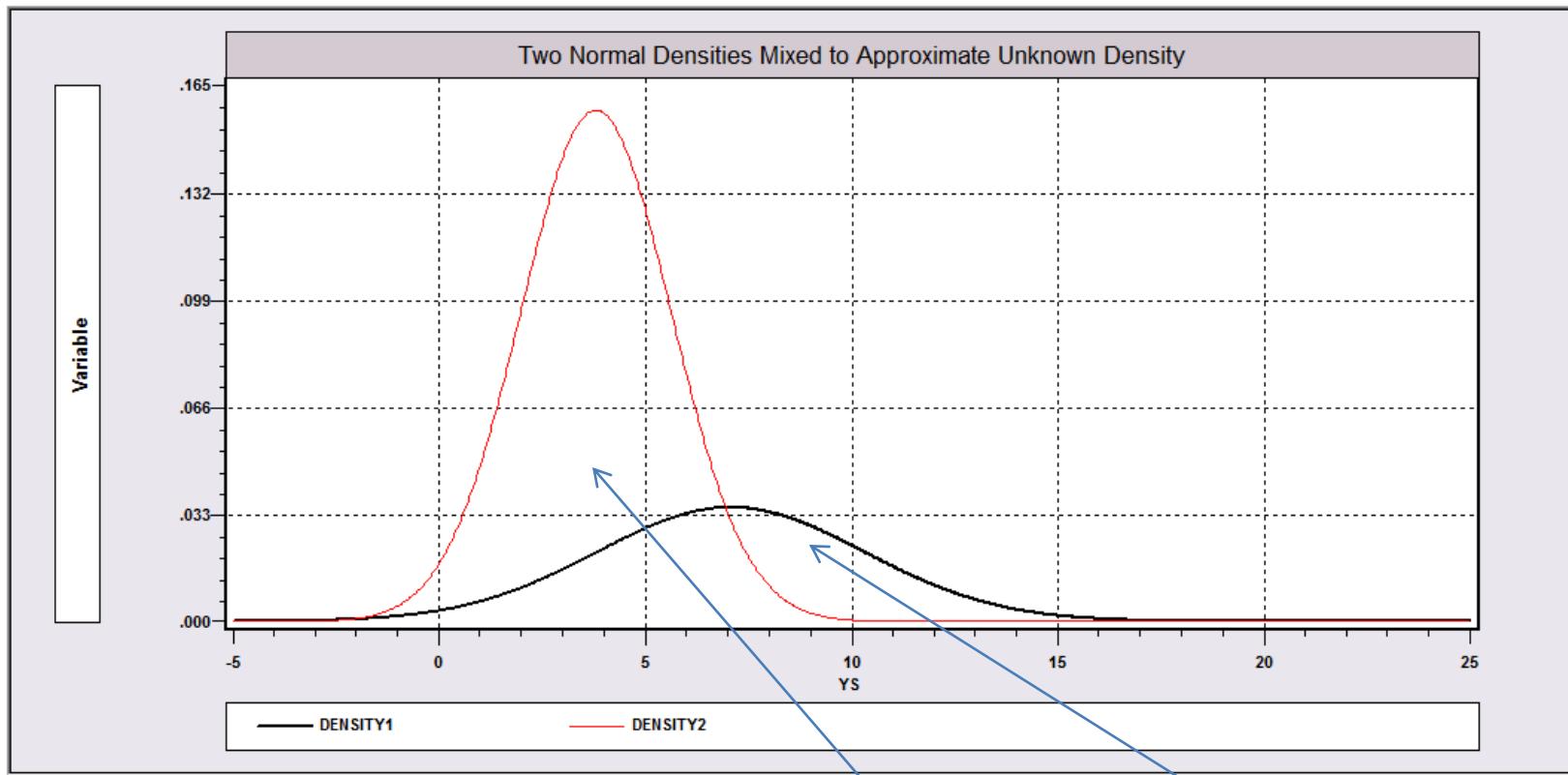
Mixture of Two Normal Densities

$$\text{LogL} = \sum_{i=1}^{1000} \log \left(\sum_{j=1}^2 \pi_j \frac{1}{\sigma_j} \phi \left(\frac{y_i - \mu_j}{\sigma_j} \right) \right)$$

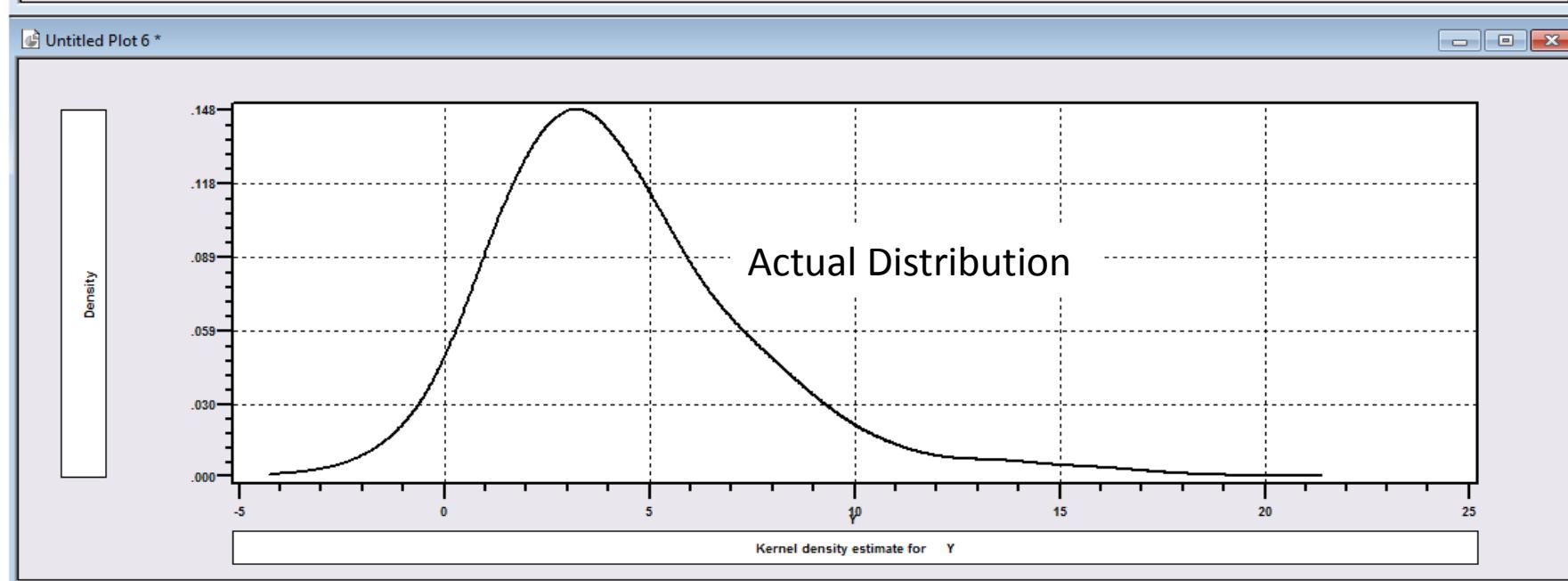
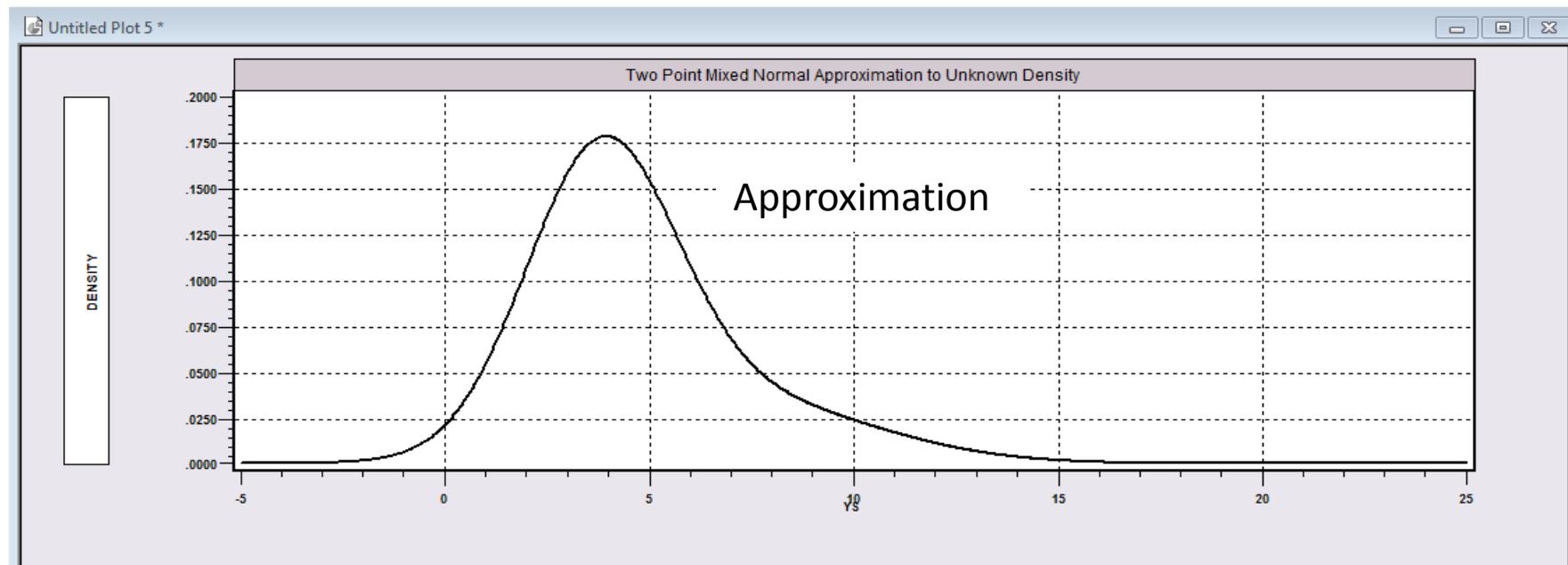
Maximum Likelihood Estimates

	Class 1		Class 2	
	Estimate	Std. Error	Estimate	Std. error
μ	7.05737	.77151	3.25966	.09824
σ	3.79628	.25395	1.81941	.10858
π	.28547	.05953	.71453	.05953

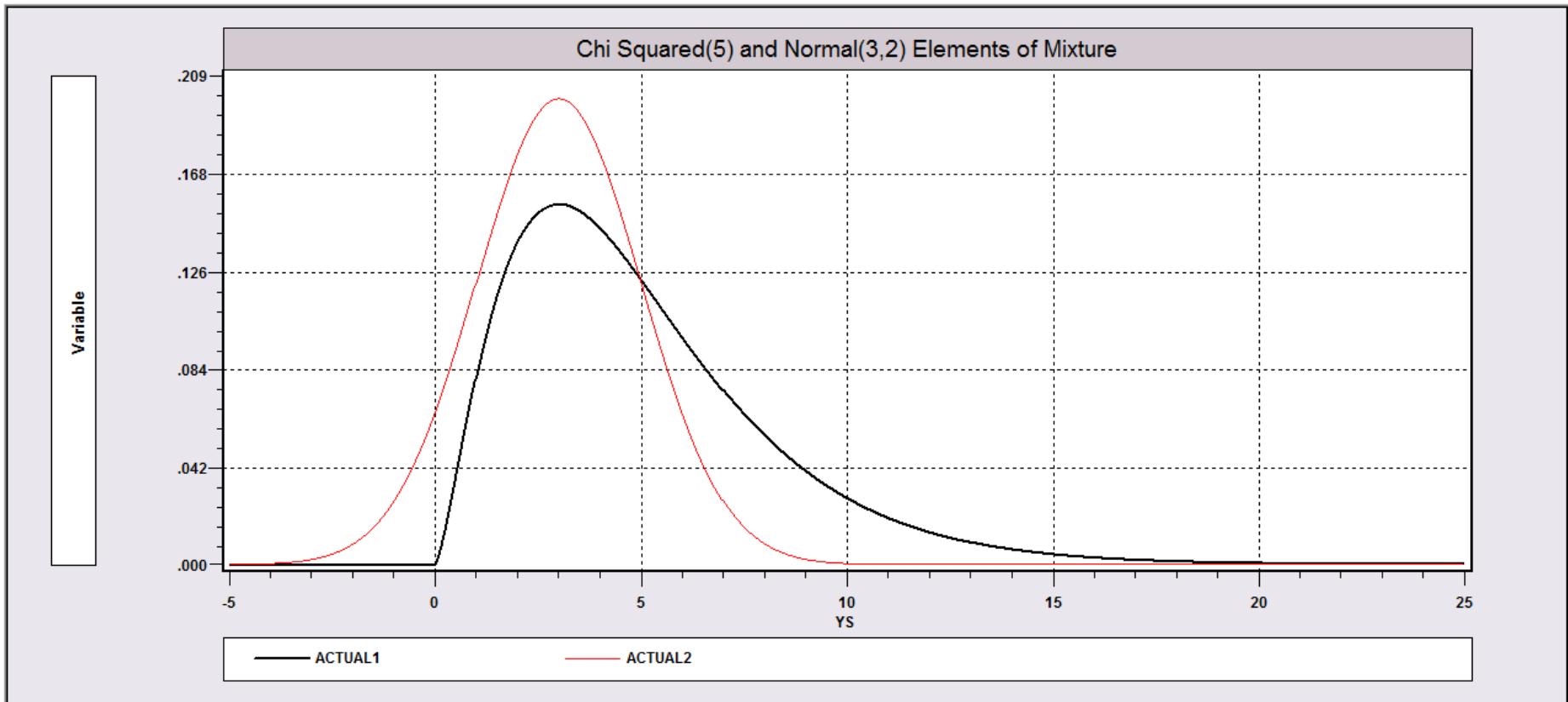
$$\hat{F}(y) = .28547 \left[\frac{1}{3.79628} \phi\left(\frac{y - 7.05737}{3.79628} \right) \right] + .71453 \left[\frac{1}{1.81941} \phi\left(\frac{y - 3.25966}{1.81941} \right) \right]$$



Mixing probabilities .715 and .285



The actual process is a mix of $\text{chi squared}(5)$ and $\text{normal}(3,2)$ with mixing probabilities .7 and .3.



$$f(y) = .7 \frac{.5^{2.5} \exp(-.5y) y^{1.5}}{\Gamma(2.5)} + .3 \frac{1}{2} \phi\left(\frac{y-3}{2}\right)$$



The Latent Class “Model”

- Parametric Model:
 - $F(y|x,\theta)$
 - E.g., $y \sim N[x'\beta, \sigma^2]$, $y \sim \text{Poisson}[\lambda = \exp(x'\beta)]$, etc.
 - Density $F(y|x,\Theta) \cong \sum_j \pi_j F(y|x,\theta_j)$,
 $\Theta = [\theta_1, \theta_2, \dots, \theta_J, \pi_1, \pi_2, \dots, \pi_J]$
 - Generating mechanism for an individual drawn at random from the mixed population is $F(y|x,\Theta)$.
 - Class probabilities relate to a stable process governing the mixture of types in the population



Latent Classes

- Population contains a mixture of individuals of different types
 - Common form of the generating mechanism within the classes
 - Observed outcome y is governed by the **common process** $F(y | x, \theta_j)$
 - Classes are distinguished by the parameters, θ_j .

**Journal of Health Economics**

Volume 28, Issue 2, March 2009, Pages 265-279

[doi:10.1016/j.jhealeco.2008.11.002](#) | How to Cite or Link Using DOI [Permissions & Reprints](#)

Health care utilisation in Europe: New evidence from the ECHP

Teresa Bago d'Uva^{a, b}, ,  and Andrew M. Jones^c

^aDepartment of Applied Economics (Room H13-09), Erasmus School of Economics, PB 1738, 3000 DR Rotterdam, The Netherlands

^bNetspar, The Netherlands

^cDepartment of Economics and Related Studies, University of York, York, YO10 5DD, United Kingdom

Received 9 July 2007; revised 30 October 2008; accepted 7 November 2008. Available online 19 November 2008.



An LC Hurdle NB2 Model

- Analysis of ECHP panel data (1994-2001)
- Two class Latent Class Model
 - Typical in health economics applications
- Hurdle model for physician visits
 - Poisson hurdle for participation and intensity given participation
 - Contrast to a negative binomial model



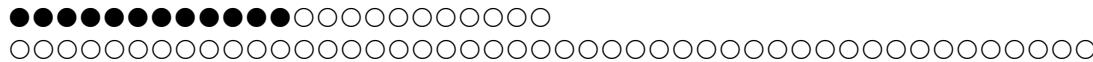
Table 8. Estimated income coefficients and elasticities for GP and specialist visits—country-specific LC hurdle models.

Country	GPs				Specialists			
	Low users		High users		Low users		High users	
	Estimated coefficient	Estimated elasticity	Estimated coefficient	Estimated elasticity	Estimated coefficient	Estimated elasticity	Estimated coefficient	Estimated elasticity
Austria	$P(Y>0)$ -0.051 (-1.467) $E(Y Y>0)$ 0.012 (0.693)	-0.012 0.009	-0.109 (-0.872) 0.039 (2.167)	-0.005 0.035	0.191 (3.743) 0.014 (0.210)	0.110 0.006	0.211 (3.556) 0.105 (3.858)	0.030 0.070
Belgium	$P(Y>0)$ 0.035 (1.002) $E(Y Y>0)$ -0.052 (-3.125)	0.008 -0.037	0.292 (4.004) -0.055 (-4.030)	0.010 -0.050	0.054 (1.399) -0.112 (-1.611)	0.036 -0.052	0.079 (1.348) -0.049 (-1.920)	0.014 -0.035
Denmark	$P(Y>0)$ 0.083 (1.746) $E(Y Y>0)$ 0.042 (0.992)	0.033 0.021	0.261 (2.302) -0.030 (-1.009)	0.023 -0.024	0.053 (0.738) -0.053 (-0.434)	0.045 -0.022	0.079 (1.123) -0.082 (-1.120)	0.034 -0.050
Finland	$P(Y>0)$ 0.054 (1.358) $E(Y Y>0)$ 0.007 (0.237)	0.024 0.004	-0.030 (-0.263) -0.048 (-1.706)	-0.003 -0.037	0.203 (3.525) -0.229 (-2.985)	0.155 -0.090	0.167 (1.909) 0.025 (0.487)	0.041 0.014
Greece	$P(Y>0)$ 0.012 (0.565) $E(Y Y>0)$ -0.024 (-1.864)	0.006 -0.015	0.015 (0.447) 0.026 (1.967)	0.004 0.020	0.184 (7.641) 0.017 (0.878)	0.128 0.010	0.148 (5.413) 0.067 (4.192)	0.060 0.055
Ireland	$P(Y>0)$ 0.164 (4.754) $E(Y Y>0)$ -0.095 (-3.865)	0.064 -0.057	0.026 (0.339) -0.049 (-2.528)	0.003 -0.043	0.172 (3.274) 0.063 (0.738)	0.152 0.027	0.313 (4.367) -0.091 (-1.838)	0.144 -0.057
Italy	$P(Y>0)$ -0.001 (-0.054) $E(Y Y>0)$ -0.044 (-4.944)	0.000 -0.031	0.116 (3.766) -0.024 (-2.691)	0.011 -0.021	0.136 (6.251) -0.084 (-2.787)	0.105 -0.035	0.190 (7.918) 0.000 (-0.026)	0.063 0.000
The Netherlands	$P(Y>0)$ 0.082 (2.897) $E(Y Y>0)$ -0.037 (-1.484)	0.035 -0.019	0.094 (1.739) -0.085 (-5.446)	0.009 -0.068	0.071 (2.085) -0.250 (-4.377)	0.055 -0.129	-0.055 (-1.084) -0.008 (-0.299)	-0.016 -0.006
Portugal	$P(Y>0)$ 0.223 (10.888) $E(Y Y>0)$ 0.027 (2.302)	0.104 0.018	0.243 (8.070) 0.001 (0.078)	0.036 0.001	0.252 (9.190) -0.087 (-3.292)	0.198 -0.045	0.295 (9.454) 0.041 (2.340)	0.099 0.028
Spain	$P(Y>0)$ -0.015 (-0.997) $E(Y Y>0)$ -0.053 (-4.401)	-0.006 -0.034	0.037 (1.261) -0.025 (-2.324)	0.005 -0.021	0.112 (5.680) -0.070 (-2.460)	0.080 -0.033	0.138 (5.189) 0.017 (1.026)	0.042 0.012

Notes: t-statistics of coefficients in parentheses. Coefficients in bold are those significant at 5%. Elasticities are calculated for each individual and averaged over the sample. Elasticities in bold correspond to significant coefficients.

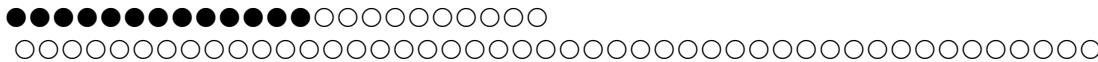
A Practical Distinction

- **Finite Mixture (Discrete Mixture):**
 - Functional form strategy
 - Component densities have no meaning
 - Mixing probabilities have no meaning
 - There is no question of “class membership”
- **Latent Class:**
 - Mixture of subpopulations
 - Component densities are believed to be definable “groups”
(Low Users and High Users)
 - The classification problem is interesting – who is in which class?
 - Posterior probabilities, $P(\text{class} | y, x)$ have meaning
 - Question of the number of classes has content in the context of the analysis



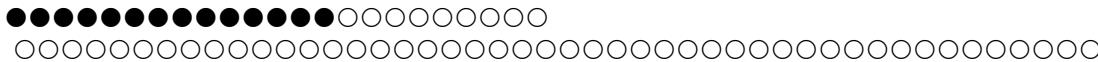
Why Make the Distinction?

- Same estimation strategy
- Same estimation results
- Extending the Latent Class Model
 - Allows a rich, flexible model specification for behavior
 - **The classes may be governed by different processes**



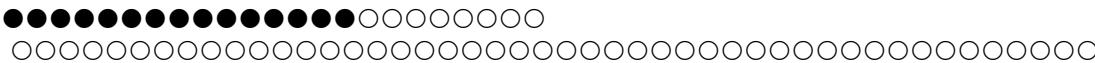
Antecedents

- Long history of finite mixture and latent class modeling in statistics and econometrics.
 - Early work starts with Pearson's 1894 study of crabs in Naples – finite mixture of two normals – looking for evidence of two subspecies.
 - See McLachlan and Peel (2000).
- Some of the extensions I will note here have already been employed in earlier literature (and noticed in surveys)
 - Different underlying processes
 - Heterogeneous class probabilities
 - Correlations of unobservables in class probabilities with unobservables in structural (within class) models
- One has not and is not widespread (yet)
 - Cross class restrictions implied by the theory of the model



Split Population Survival Models

- Schmidt and Witte 1989 study of recidivism
- **F=1 for eventual failure, F=0 for never fail.** Unobserved.
 $P(F=1)=d$, $P(F=0)=1-d$
- $C=1$ for recidivist, observed. $\text{Prob}(F=1 | C=1)=1$.
- Density for time until failure actually occurs is $d \times g(t | F=1)$.
- Density for observed duration (possibly censored)
 - $P(C=0)=(1-d)+d(G(T | F=1))$ (Observation is censored)
 - Density given $C=1 = dg(t | F=1)$
 - G =survival function, t =time of observation.
- **Unobserved F implies a latent population split.**
- They added covariates to d : $d_i = \text{logit}(z_i)$.
- **Different models apply to the two latent subpopulations.**

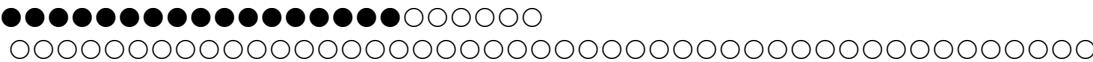


Switching Regressions

- Mixture of normals with heterogeneous mean
 - $y \sim N(\mathbf{b}_0' \mathbf{x}, \sigma_0^2)$ if $d=0$, $y \sim N(\mathbf{b}_1' \mathbf{x}, \sigma_1^2)$ if $d=1$
 - d is unobserved (Latent switching). $P(d=1) = f(\mathbf{c}' \mathbf{z})$.
 - Lack of identification (Kiefer)
 - Becomes a latent class model when regime 0 is a demand function and regime 1 is a supply function, $d=0$ if excess supply
 - **The two regression equations may involve different variables – a true latent class model**

Applications

- The Union-nonunion-wage model (Lee, 1978)
- The Housing-demand model (Trost, 1977)
- Disequilibrium Market model (Fair and Jaffee, 1972)
- The Labor-supply model (Heckman, 1974)
- The Labor-supply model (Gronau, 1974)
- Needs vs. Reluctance model (Polakoff and Sibler, 1967)



Endogenous Switching

$$\text{Regime 0: } y_i = x'_{i0} \beta_0 + \varepsilon_{i0}$$

$$\text{Regime 1: } y_i = x'_{i1} \beta_1 + \varepsilon_{i1}$$

$$\text{Regime Switch: } d^* = z'_i \gamma + u, \quad d = 1[d^* > 0]$$

Regime 0 governs if $d = 0$, Probability = $1 - \Phi(z'_i \gamma)$

Regime 1 governs if $d = 1$, Probability = $\Phi(z'_i \gamma)$

Not identified. Regimes do not coexist.

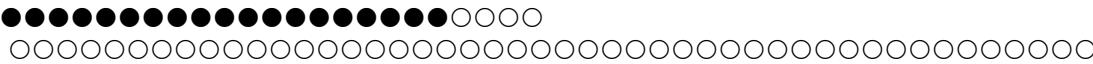
$$\text{Endogenous Switching: } \begin{pmatrix} \varepsilon_{i0} \\ \varepsilon_{i0} \\ \varepsilon_{i0} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & ? & ? \\ ? & \sigma_1^2 & ? \\ ? & ? & 1 \end{pmatrix} \right]$$

This is a latent class model with different processes in the two classes.

There is correlation between the unobservables that govern the class determination and the unobservables in the two regime equations.

Zero Inflation Models

- Lambert 1992, Technometrics. Quality control problem. Counting defects per unit of time on the assembly line. How to explain the zeros; is the process under control or not?
- Two State Outcome: $\text{Prob}(\text{State}=0)=R$, $\text{Prob}(\text{State}=1)=1-R$
 - State=0, $Y=0$ with certainty
 - State=1, $Y \sim \text{some distribution support that includes } 0$, e.g., Poisson.
- $\text{Prob}(\text{State } 0 | y>0) = 0$
- $\text{Prob}(\text{State } 1 | y=0) = (1-R)f(0)/[R + (1-R)f(0)]$
- R = Logistic probability (with the same covariates as f – the “zip-tau” model)
- “Nonstandard” latent class model according to McLachlan and Peel
- Recent users have extended this to “Outcome Inflated Models,” e.g., two inflation in models of fertility.



Variations of Interest

- Heterogeneous priors for the class probabilities
- Correlation of unobservables in class probabilities with unobservables in regime specific models
- Variations of model structure across classes
- Behavioral basis for the mixed model with implied restrictions

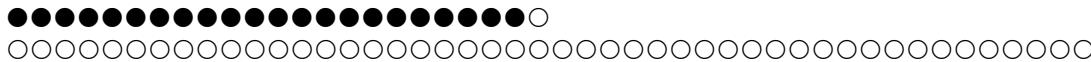


Heterogeneous Class Probabilities

- $\pi_j = \text{Prob}(\text{class}=j) = \text{governor of a detached natural process.}$
 - $\pi_{ij} = \text{Prob}(\text{class}=j | \text{individual } i)$
 - Usually modeled as a multinomial logit
 - Now possibly a behavioral aspect of the process, no longer “detached” or “natural”
 - $F(y_i | x_i, \Theta) \cong \sum_j \pi_{ij}(z_i, \delta) F(y_i | x_i, \theta_j)$
 - Nagin and Land 1993, “Criminal Careers...

Interpreting the Discrete Variation

Most empirical applications of latent class models to health care utilisation take class membership probabilities as parameters $\pi_{ij} = \pi_j, j=1, \dots, C$ to be estimated along with $\vartheta_1, \dots, \vartheta_C$ (e.g., P. Deb and P.K. Trivedi, Demand for medical care by the elderly: a finite mixture approach, *Journal of Applied Econometrics* **12** (1997), pp. 313–336 [[Deb and Trivedi, 1997](#)], [\[Deb, 2001\]](#), [\[Jiménez-Martín et al., 2002\]](#), [\[Atella et al., 2004\]](#) and [\[Bago d'Uva, 2006\]](#)). This is analogous to the hypothesis that individual heterogeneity is uncorrelated with the regressors in a random effects or random parameters specification. A more general approach is to parameterise the heterogeneity as a function of time invariant individual characteristics z_i , as in [Mundlak \(1978\)](#), thus accounting for the possible correlation between observed regressors and unobserved effects. This has been done in recent studies that consider continuous distributions for the individual effects, mostly by setting $z_i = \bar{x}_i$. To implement this approach in the case of the latent class model, class membership can be modelled as a multinomial logit (as in, e.g., [\[Clark and Etilé, 2006\]](#), [\[Clark et al., 2005\]](#) and [\[Bago d'Uva, 2005\]](#)):



A Loose End in the Theory

Accounting for correlation between regressors and unobserved effects using the Mundlak approach in a regression model:

Uncorrelated $y_{it} = \mathbf{x}'_{it}\beta + \alpha_i + \varepsilon_{it},$
 $\alpha_i = \alpha + u_i, \quad u_i \perp \mathbf{x}_{it}, \quad E[u_i | \mathbf{x}_{it}] = 0$

Correlated $y_{it} = \mathbf{x}'_{it}\beta + \alpha_i + \varepsilon_{it},$
 $\alpha_i = \alpha + u_i, \quad u_i \text{ not } \perp \mathbf{x}_{it}, \quad E[u_i | \mathbf{x}_{it}] \neq 0$

Mundlak correction

$$\alpha_i = \bar{\mathbf{x}}_i \delta + u_i, \quad E[u_i | \mathbf{x}_{it}, \bar{\mathbf{x}}_i] \neq 0$$

A Mundlak Correction for LCM?

The latent class model with heterogeneous priors

Uncorrelated: $F(y_{it} | \mathbf{x}_{it}, \text{class} = j) = F(y_{it} | \mathbf{x}_{it}, \boldsymbol{\theta}_j)$

$$\boldsymbol{\theta}_j = \boldsymbol{\theta} + \Delta_j,$$

$$E[\Delta_j] = 0 \text{ over discrete support } \Delta_j = \Delta_1, \Delta_2, \dots, \Delta_J,$$

$$\text{Prob}(\Delta_j) = \pi_j, j = 1, \dots, J \text{ independent of } \mathbf{z}_i$$

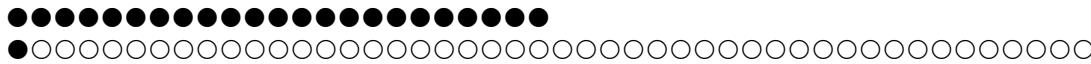
Correlated: $F(y_{it} | \mathbf{x}_{it}, \text{class} = j) = F(y_{it} | \mathbf{x}_{it}, \boldsymbol{\theta}_j)$

$$\boldsymbol{\theta}_j = \boldsymbol{\theta} + \Delta_j,$$

$$E[\Delta_j] = 0 \text{ over discrete support } \Delta_j = \Delta_1, \Delta_2, \dots, \Delta_J,$$

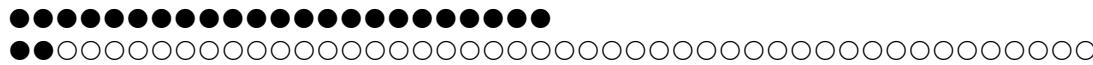
$$\text{Prob}(\Delta_j | \mathbf{z}_i) = \pi_{ij} = \frac{\exp(\mathbf{z}'_i \boldsymbol{\delta}_j)}{\sum_{j=1}^J \exp(\mathbf{z}'_i \boldsymbol{\delta}_j)}, j = 1, \dots, J$$

- (1) This does not make Δ_j uncorrelated with \mathbf{x}_{it}
- (2) It makes $\text{Prob}(\Delta_j)$ correlated with \mathbf{x}_{it} . They may have been already. It is not clear from the model specification.
- (3) To do the equivalent of Mundlak, we would need $\Delta_{ij} = f(\mathbf{z}_i) + \mathbf{h}_{ij}$



Applications

- Obesity: Correlation of Unobservables
- Self Assessed Health: Heterogeneous subpopulations
- Choice Strategy in Travel Route Choice: Cross Class Restrictions
- Cost Efficiency of Nursing Homes: Theoretical Restrictions on Underlying Models
- Freight Forwarding: Finite Mixture of Random Parameters Models



A Bivariate Latent Class Correlated Generalized Ordered Probit Model with an Application to Modeling Observed Obesity Levels*

William Greene^a, Mark N. Harris^b, Bruce Hollingsworth^c
and Pushkar Maitra^d

^aEconomics Department, Stern Business School,
New York University;

^bDepartment of Economics and Business Statistics,
Monash University;

^cCentre for Health Economics, Monash University;

^dDepartment of Economics, Monash University.



Modeling BMI

WHO BMI Classes: 1 – 4

Standard ordered probit model

$$y = \begin{cases} 0 & \text{if } y^* \leq 0 \\ j & \text{if } \mu_{j-1} < y^* \leq \mu_j, \text{ for } 0 < j < J \\ J & \text{if } \mu_{J-1} \leq y^*, \end{cases}$$

$$\Pr(y) = \begin{cases} \Pr(y = 0 | z) &= \Phi(\mu_0 - \mathbf{z}' \boldsymbol{\gamma}) \\ \Pr(y = j | z) &= \Phi(\mu_j - \mathbf{z}' \boldsymbol{\gamma}) - \Phi(\mu_{j-1} - \mathbf{z}' \boldsymbol{\gamma}); \text{ for } 1 < j < J \\ \Pr(y = J | z) &= 1 - \Phi(\mu_{J-1} - \mathbf{z}' \boldsymbol{\gamma}) \end{cases}$$



Standard Two Class LCM

$$c^* = \mathbf{x}' \boldsymbol{\beta} + \varepsilon$$

$$\Pr(c=1|\mathbf{x}) = \Pr(c^* > 0|\mathbf{x}) = \Phi(\mathbf{x}' \boldsymbol{\beta}),$$

$$\begin{aligned}\Pr(y=j|\mathbf{x}, \mathbf{z}) = & \Pr(c=0|\mathbf{x}) \Pr(y=j|\mathbf{z}, c=0) \\ & + \Pr(c=1|\mathbf{x}) \Pr(y=j|\mathbf{z}, c=1).\end{aligned}$$

So, for those belonging to class 0 we have

$$\Pr(y=0|\mathbf{z}, \mathbf{x}, c=0) = (1 - \Phi(\mathbf{x}' \boldsymbol{\beta})) [\Phi(-\mathbf{z}' \boldsymbol{\gamma}_0)]$$

$$\Pr(y=j|c=0) = \Pr(y=j|\mathbf{z}, \mathbf{x}, c=0) = (1 - \Phi(\mathbf{x}' \boldsymbol{\beta})) [\Phi(\mu_{0,j} - \mathbf{z}' \boldsymbol{\gamma}_0) - \Phi(\mu_{0,j-1} - \mathbf{z}' \boldsymbol{\gamma}_0)] \quad (1 < j < J)$$

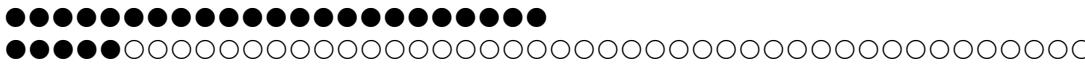
$$\Pr(y=J|\mathbf{z}, \mathbf{x}, c=0) = (1 - \Phi(\mathbf{x}' \boldsymbol{\beta})) [1 - \Phi(\mu_{0,J-1} - \mathbf{z}' \boldsymbol{\gamma}_0)],$$

And similarly for those belonging to class 1 we have

$$\Pr(y=0|\mathbf{z}, \mathbf{x}, c=1) = \Phi(\mathbf{x}' \boldsymbol{\beta}) [\Phi(-\mathbf{z}' \boldsymbol{\gamma}_1)]$$

$$\Pr(y=j|c=1) = \Pr(y=j|\mathbf{z}, \mathbf{x}, c=1) = \Phi(\mathbf{x}' \boldsymbol{\beta}) [\Phi(\mu_{1,j} - \mathbf{z}' \boldsymbol{\gamma}_1) - \Phi(\mu_{1,j-1} - \mathbf{z}' \boldsymbol{\gamma}_1)] \quad (1 < j < J)$$

$$\Pr(y=J|\mathbf{z}, \mathbf{x}, c=1) = \Phi(\mathbf{x}' \boldsymbol{\beta}) [1 - \Phi(\mu_{1,J-1} - \mathbf{z}' \boldsymbol{\gamma}_1)],$$



Correlation of Unobservables

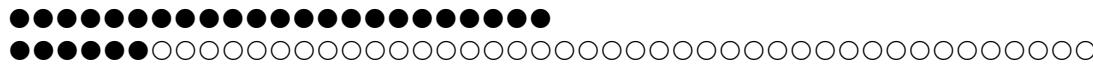
Class Probabilities

$$c^* = \mathbf{x}' \boldsymbol{\beta} + \varepsilon, \quad \Pr(c=1|\mathbf{x}) = \Pr(c^* > 0|\mathbf{x}) = \Phi(\mathbf{x}' \boldsymbol{\beta}),$$

Within Class Model (Process)

$$\mathbf{y}^* = \mathbf{z}' \boldsymbol{\gamma} + u$$

$$y = \begin{cases} 0 & \text{if } y^* \leq 0 \\ j & \text{if } \mu_{j-1} < y^* \leq \mu_j, \text{ for } 0 < j < J \\ J & \text{if } \mu_{J-1} \leq y^*, \end{cases}$$



$$\Pr(y = j, c = 1) = \begin{cases} \Pr(y = 0, c=1|x, z) &= \Phi_2(x'\beta, -z'\gamma_1; \rho_1) \\ \Pr(y = j, c=1|x, z) &= \Phi_2(x'\beta, \mu_{1,j} - z'\gamma_1; \rho_1) \\ &\quad - \Phi_2(x'\beta, \mu_{1,j-1} - z'\gamma_1; \rho_1) \quad (0 < j < J) \\ \Pr(y = J, c=1|x, z) &= \Phi_2(x'\beta, z'\gamma_1 - \mu_{1,J-1}; \rho_1) \end{cases}$$

where $\Phi_2(a,b;\rho)$ denotes the cumulative distribution function of the standardized bivariate normal distribution with correlation coefficient ρ between the univariate random elements, while those for class 0 are

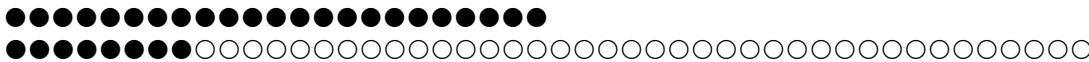
$$\Pr(y = j, c = 0) = \begin{cases} \Pr(y = 0, c=0|x, z) &= \Phi_2(-x'\beta, -z'\gamma_0; -\rho_0) \\ \Pr(y = j, c=0|x, z) &= \Phi_2(-x'\beta, \mu_{0,j} - z'\gamma_0; -\rho_0) \\ &\quad - \Phi_2(-x'\beta, \mu_{0,j-1} - z'\gamma_0; -\rho_0) \quad (0 < j < J) \\ \Pr(y = J, c=0|x, z) &= \Phi_2(-x'\beta, z'\gamma_0 - \mu_{0,J-1}; -\rho_0) \end{cases}$$

These assume c is observed. But, the right terms would be $\Pr(y=j|c=1)\Pr(c=1)$ which is, trivially, $[\Pr(y=j,c=1)/\Pr(c=1)] \times \Pr(c=1)$ which returns the preceding.



An Identification Issue in Generalized Ordered Choice Models

- $\text{Prob}(y=j | c=0) = \Phi_2(-x'\beta, \mu_{0,j} - z'\gamma_0; -\rho_0)$
 $\quad - \Phi_2(-x'\beta, \mu_{0,j-1} - z'\gamma_0; -\rho_0)$
- If $\mu_{0,j} = \delta_0 + w_i'\delta_0$ then δ_0 and γ_0 are not separately identified.
- We used $\mu_{cij} = \exp(\delta_{cj} + w_i'\delta_c)$,
- Isn't this "identification by functional form?"



Inflated Responses in Self-Assessed Health

Mark Harris

Department of Econometrics and Business Statistics, Monash University
&

William Greene

Stern Business School, NYU, New York
&

Bruce Hollingsworth

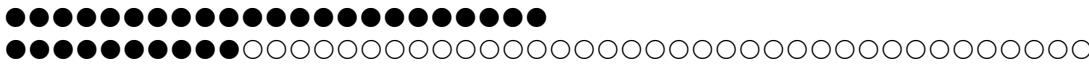
Centre for Health Economics, Monash University

FUNDING PROVIDED BY
Australian Research Council (ARC)



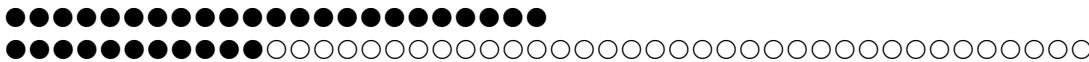
Introduction

- The typical question would be: “*In general, would you say your health is: Excellent, Very good, Good, Fair or Poor?*”
- So here respondents “tick a box”, typically from 1 – 5, for these responses
- What we typically find is that approx. $\frac{3}{4}$ of the nation are of “good” or “very good” health
 - in our data (HILDA) we get 72%
- Get similar numbers for most developed countries
- So, key question is, *does this truly represent the health of the nation?*



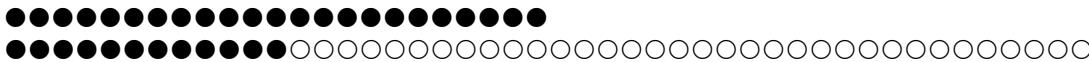
An Ordered Choice Model for Self Assessed Health

$$y = \begin{cases} 0 & \text{if } y^* \leq \mu_0 \\ 1 & \text{if } \mu_0 < y^* \leq \mu_1 \\ 2 & \text{if } \mu_1 < y^* \leq \mu_2 \\ \vdots & \vdots \\ J & \text{if } \mu_{J-1} < y^* \end{cases}$$



An Ordered Probit Model

- So, we would typically label the outcomes of this likert scale as:
 - $y = 0$: poor health
 - $y = 1$: fair health
 - $y = 2$: good health
 - $y = 3$: very good health
 - $y = 4$: excellent health
 - (or similar)
- And no matter the exact wording, it's always outcomes 2 and 3 that clearly dominate



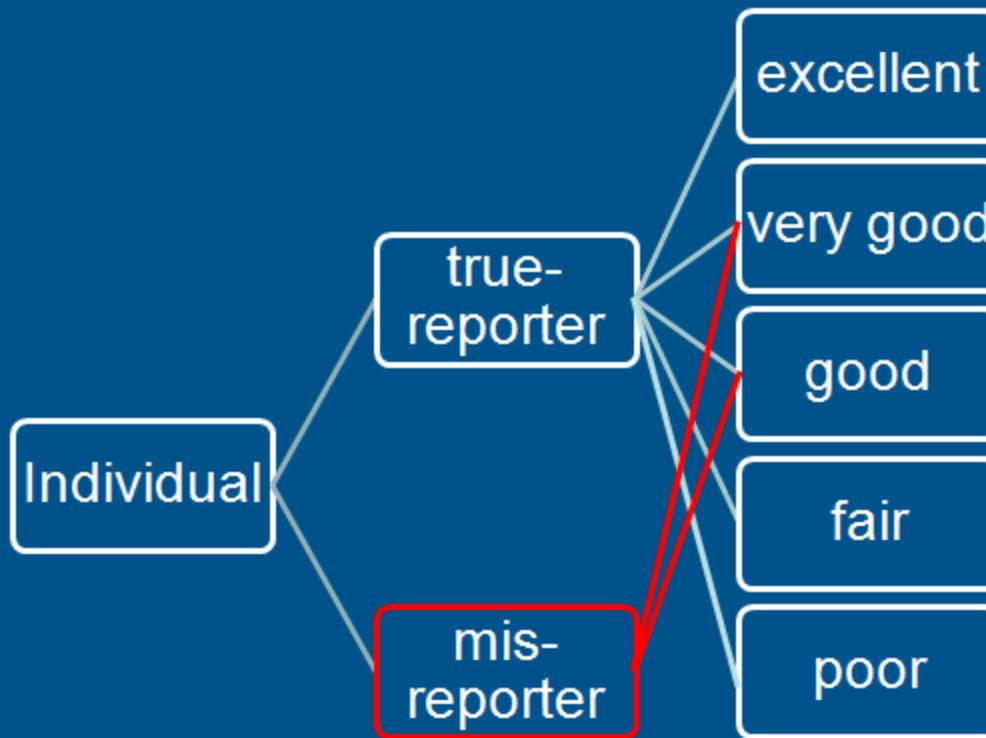
A Multiple Inflated Ordered Probit Model

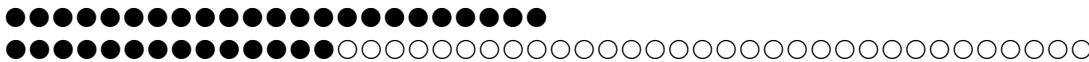
- True-reporters can choose any outcome
- This actual response will be determined by a further latent variable equation, along the lines of the previous Ordered Probit set-up...

$$y^* = x'_y \beta_y + \varepsilon_y$$



A Multiple Inflated Ordered Probit Model





A Multiple Inflated Ordered Probit Model

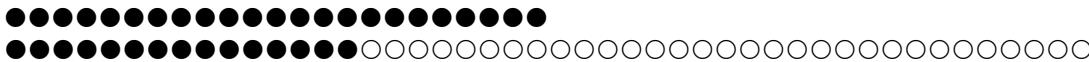
- If we maintain the assumption of normality (and independence), the probability of a mis-reporter is simply a Probit probability:

$$\Pr(r^* < 0) = \Phi(-x_r' \beta_r)$$

- And the joint probability of a mis-reporter reporting very good and good, respectively, are products of Probit probabilities:

$$\Pr(\text{mis, very good}) = \Phi(-x_r' \beta_r) \Phi(x_m' \beta_m)$$

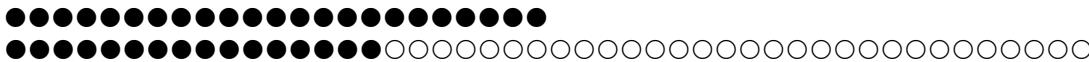
$$\Pr(\text{mis, good}) = \Phi(-x_r' \beta_r) \Phi(-x_m' \beta_m)$$



A Multiple Inflated Ordered Probit Model

- For true-reporters the respective probabilities for each outcome are simply the ordered probit ones, weighted by the probability of being a truthful respondent (0=poor to 4=excellent):

$$\Pr(true, y) = \begin{cases} 0 = \Phi(x'_r \beta_r) \times [\Phi(-x'_y \beta_y)] \\ 1 = \Phi(x'_r \beta_r) \times [\Phi(\mu_1 - x'_y \beta_y) - \Phi(-x'_y \beta_y)] \\ 2 = \Phi(x'_r \beta_r) \times [\Phi(\mu_2 - x'_y \beta_y) - \Phi(\mu_1 - x'_y \beta_y)] \\ 3 = \Phi(x'_r \beta_r) \times [\Phi(\mu_3 - x'_y \beta_y) - \Phi(\mu_2 - x'_y \beta_y)] \\ 4 = \Phi(x'_r \beta_r) \times [1 - \Phi(\mu_3 - x'_y \beta_y)] \end{cases}$$

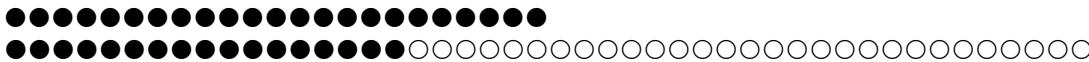


A Multiple Inflated Ordered Probit Model

- Putting all these elements together we have a latent class model:

$$\Pr(y) = \begin{cases} 0 = \Phi(x'_r \beta_r) \times [\Phi(-x'_y \beta_y)] \\ 1 = \Phi(x'_r \beta_r) \times [\Phi(\mu_1 - x'_y \beta_y) - \Phi(-x'_y \beta_y)] \\ 2 = \Phi(x'_r \beta_r) \times [\Phi(\mu_2 - x'_y \beta_y) - \Phi(\mu_1 - x'_y \beta_y)] + \Phi(-x'_r \beta_r) \Phi(-x'_m \beta_m) \\ 3 = \Phi(x'_r \beta_r) \times [\Phi(\mu_3 - x'_y \beta_y) - \Phi(\mu_2 - x'_y \beta_y)] + \Phi(-x'_r \beta_r) \Phi(x'_m \beta_m) \\ 4 = \Phi(x'_r \beta_r) \times [1 - \Phi(\mu_3 - x'_y \beta_y)] \end{cases}$$

- So this multiple inflated ordered probit model gives the good and very good categories an additional

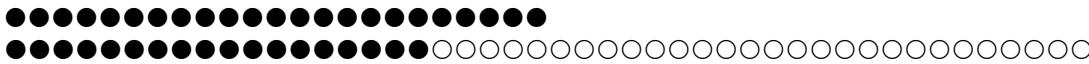


A Multiple Inflated Ordered Probit Model

- With nonzero correlations, probabilities are now functions of bivariate normal cdf's:

$$\begin{cases} 0 = \Phi_2(x_r'\beta_r, -x_y'\beta_y; -\rho_{ny}) \\ 1 = \Phi_2(x_r'\beta_r, \mu_1 - x_y'\beta_y; -\rho_{ny}) - \Phi_2(x_r'\beta_r, -x_y'\beta_y; -\rho_{ny}) \\ 2 = [\Phi_2(x_r'\beta_r, \mu_2 - x_y'\beta_y; -\rho_{ny}) - \Phi_2(x_r'\beta_r, \mu_1 - x_y'\beta_y; -\rho_{ny})] + \Phi_2(-x_r'\beta_r, -x_m'\beta_m; \rho_{nm}) \\ 3 = [\Phi_2(x_r'\beta_r, \mu_3 - x_y'\beta_y; -\rho_{ny}) - \Phi_2(x_r'\beta_r, \mu_2 - x_y'\beta_y; -\rho_{ny})] + \Phi_2(-x_r'\beta_r, x_m'\beta_m; -\rho_{nm}) \\ 4 = \Phi_2(x_r'\beta_r, x_y'\beta_y - \mu_3; \rho_{ny}) \end{cases}$$

- Again inflates the good/very good categories



Choice Strategy

Hensher, D.A., Rose, J. and Greene, W. (2005) The Implications on Willingness to Pay of **Respondents Ignoring Specific Attributes** (DoD#6) *Transportation*, 32 (3), 203-222.

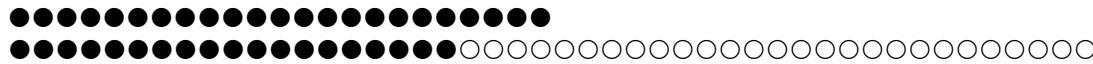
Hensher, D.A. and Rose, J.M. (2009) **Simplifying Choice through Attribute Preservation or Non-Attendance**: Implications for Willingness to Pay, *Transportation Research Part E*, 45, 583-590.

Rose, J., Hensher, D., Greene, W. and Washington, S. **Attribute Exclusion Strategies** in Airline Choice: Accounting for Exogenous Information on Decision Maker Processing Strategies in Models of Discrete Choice, *Transportmetrica*, 2011

Hensher, D.A. and Greene, W.H. (2010) **Non-attendance and dual processing of common-metric attributes** in choice analysis: a latent class specification, *Empirical Economics* 39 (2), 413-426

Campbell, D., Hensher, D.A. and Scarpa, **R. Non-attendance to Attributes in Environmental Choice Analysis**: A Latent Class Specification, *Journal of Environmental Planning and Management*, proofs 14 May 2011.

Hensher, D.A., Rose, J.M. and Greene, W.H. **Inferring attribute non-attendance from stated choice data**: implications for willingness to pay estimates and a warning for stated choice experiment design, 14 February 2011, *Transportation*, online 2 June 2001 DOI 10.1007/s11116-011-9347-8.



Decision Strategy in Multinomial Choice

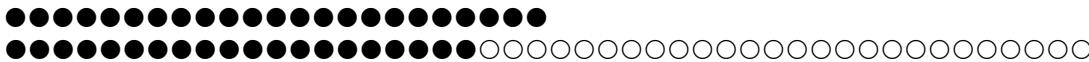
Choice Situation: Alternatives A_1, \dots, A_J

Attributes of the choices: x_1, \dots, x_K

Characteristics of the individual: z_1, \dots, z_M

Random utility functions: $U(j|x, z) = U(x_{ij}, z_j, \varepsilon_{ij})$

Choice probability model: $\text{Prob}(\text{choice}=j) = \text{Prob}(U_j > U_l) \quad \forall l \neq j$



Stated Choice Experiment

Transport Study

Games 1

	Details of Your Recent Trip	Alternative Road A	Alternative Road B	Alternative Road C
Time in free-flow (mins)	15	14	16	16
Time slowed down by other traffic (mins)	10	12	8	12
Time in Stop/Start conditions (mins)	5	4	6	4
Uncertainty in travel time (mins)	+/- 10	+/- 12	+/- 8	+/- 8
Running costs	\$ 2.20	\$ 2.40	\$ 2.40	\$ 2.10
Toll costs	\$ 2.00	\$ 2.10	\$ 2.10	\$ 1.90

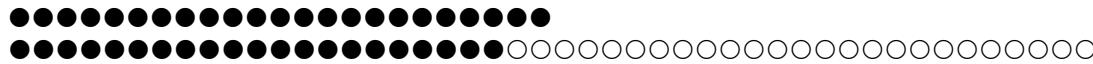
If you take the same trip again, which road would you choose?

Current Road Road A Road B Road C

If you could only choose between the new roads, which would you choose?

Road A Road B Road C

[Go to Game 2 of 6](#)

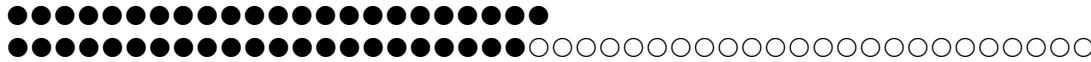


Multinomial Logit Model

$$\text{Prob}(\text{choice} = j) = \frac{\exp[\beta' x_{ij} + \gamma'_j z_i]}{\sum_{j=1}^J \exp[\beta' x_{ij} + \gamma'_j z_i]}$$

Behavioral model assumes

- (1) Utility maximization (and the underlying micro- theory)
- (2) Individual pays attention to all attributes. That is the implication of the nonzero β .



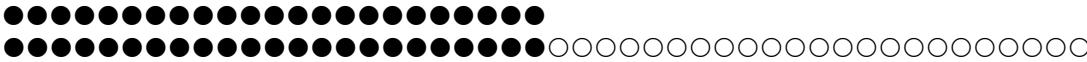
Individual Explicitly Ignores Attributes

Hensher, D.A., Rose, J. and Greene, W. (2005) The Implications on Willingness to Pay of Respondents Ignoring Specific Attributes (DoD#6) *Transportation*, 32 (3), 203-222.

Hensher, D.A. and Rose, J.M. (2009) Simplifying Choice through Attribute Preservation or Non-Attendance: Implications for Willingness to Pay, *Transportation Research Part E*, 45, 583-590.

Rose, J., Hensher, D., Greene, W. and Washington, S. Attribute Exclusion Strategies in Airline Choice: Accounting for Exogenous Information on Decision Maker Processing Strategies in Models of Discrete Choice, *Transportmetrica*, 2011

Choice situations in which the individual explicitly states that they ignored certain attributes in their decisions.



Stated Choice Experiment

Transport Study

Games 1

	Details of Your Recent Trip	Alternative Road A	Alternative Road B	Alternative Road C
Time in free-flow (mins)	15	14	16	16
Time slowed down by other traffic (mins)	10	12	8	12
Time in Stop/Start conditions (mins)	5	4	6	4
Uncertainty in travel time (mins)	+/- 10	+/- 12	+/- 8	+/- 8
Running costs	\$ 2.20	\$ 2.40	\$ 2.40	\$ 2.10
Toll costs	\$ 2.00	\$ 2.10	\$ 2.10	\$ 1.90

If you take the same trip again, which road would you choose?

Current Road Road A Road B Road C

If you could only choose between the new roads, which would you choose?

Road A Road B Road C

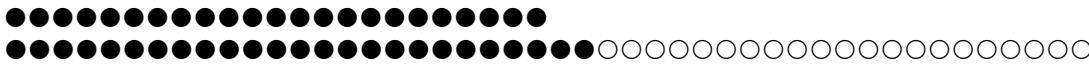
[Go to Game 2 of 6](#)

Ancillary questions: Did you ignore any of these attributes?



Appropriate Modeling Strategy

- Fix ignored attributes at zero? Definitely not!
 - Zero is an unrealistic value of the attribute (price)
 - The probability is a function of $x_{ij} - x_{il}$, so the substitution distorts the probabilities
- Appropriate model: for that individual, the specific coefficient is zero – consistent with the utility assumption. A person specific, exogenously determined model
- Surprisingly simple to implement

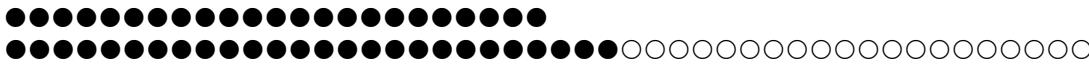


Individual Implicitly Ignores Attributes

Hensher, D.A. and Greene, W.H. (2010) Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification, *Empirical Economics* 39 (2), 413-426

Campbell, D., Hensher, D.A. and Scarpa, R. Non-attendance to Attributes in Environmental Choice Analysis: A Latent Class Specification, *Journal of Environmental Planning and Management*, proofs 14 May 2011.

Hensher, D.A., Rose, J.M. and Greene, W.H. Inferring attribute non-attendance from stated choice data: implications for willingness to pay estimates and a warning for stated choice experiment design, 14 February 2011, *Transportation*, online 2 June 2001 DOI 10.1007/s11116-011-9347-8.



Stated Choice Experiment

Transport Study

Games 1

	Details of Your Recent Trip	Alternative Road A	Alternative Road B	Alternative Road C
Time in free-flow (mins)	15	14	16	16
Time slowed down by other traffic (mins)	10	12	8	12
Time in Stop/Start conditions (mins)	5	4	6	4
Uncertainty in travel time (mins)	+/- 10	+/- 12	+/- 8	+/- 8
Running costs	\$ 2.20	\$ 2.40	\$ 2.40	\$ 2.10
Toll costs	\$ 2.00	\$ 2.10	\$ 2.10	\$ 1.90

If you take the same trip again, which road would you choose?

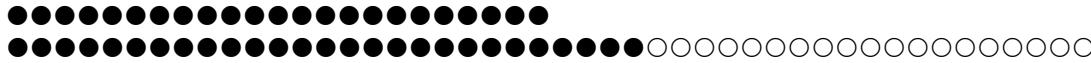
Current Road Road A Road B Road C

If you could only choose between the new roads, which would you choose?

Road A Road B Road C

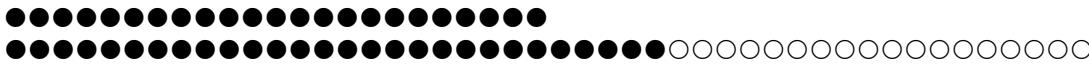
[Go to Game 2 of 6](#)

Individuals seem to be ignoring attributes. Unknown to the analyst



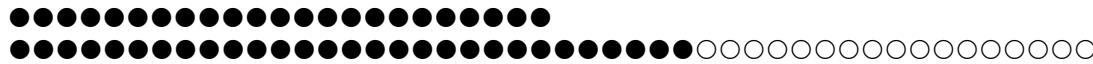
The 2^K model

- The analyst believes some attributes are ignored. There is no indicator.
- Classes distinguished by which attributes are ignored
- Same model applies now a latent class. For K attributes there are 2^K candidate coefficient vectors

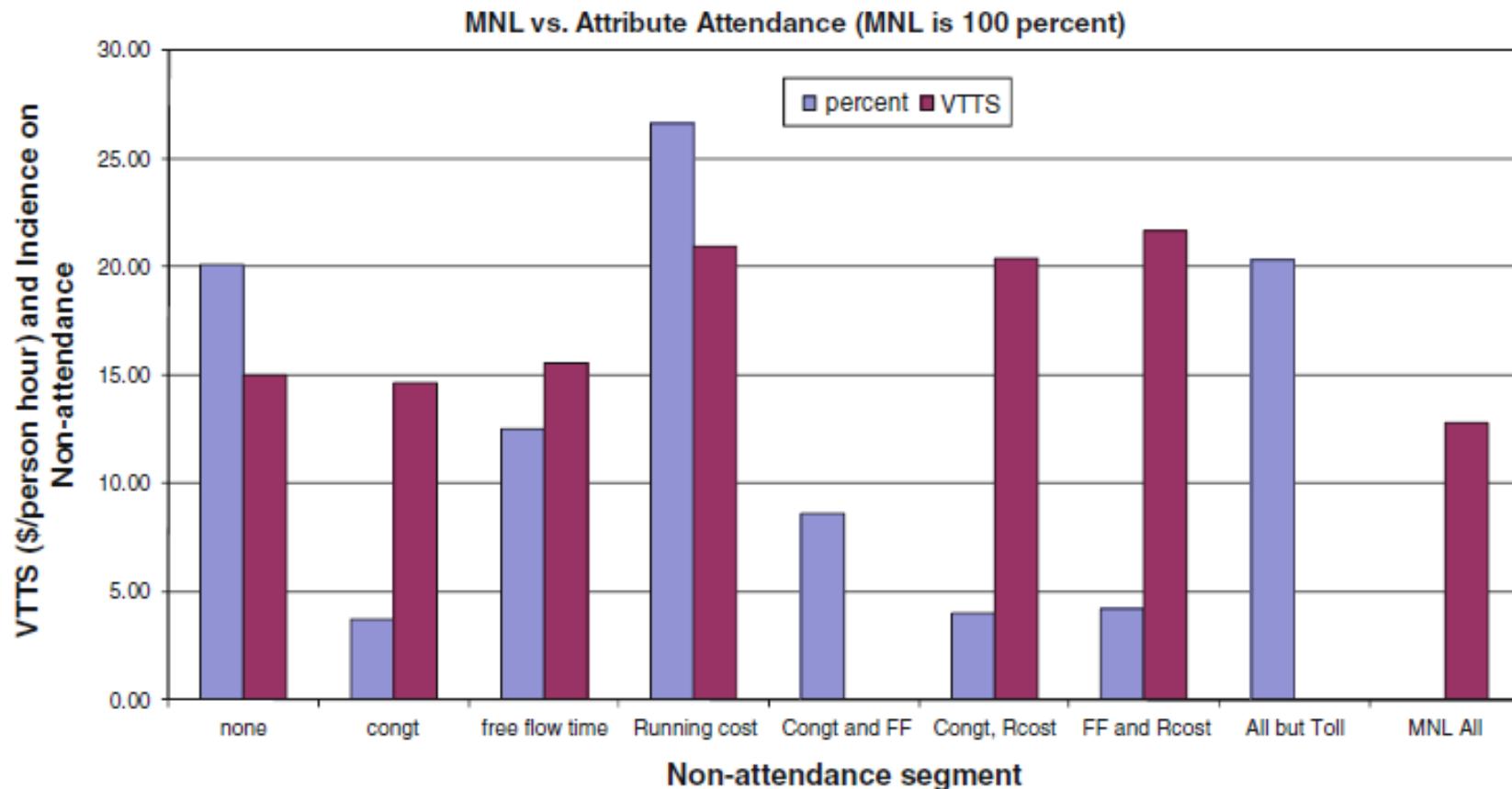


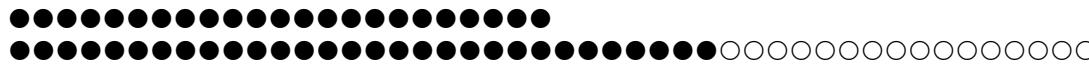
A Latent Class Model

$$\left\{ \begin{array}{c} \left[\begin{array}{ccc} \text{Uncertainty} & \text{Toll Cost} & \text{Running Cost} \\ \beta_1 & \beta_2 & \beta_3 \end{array} \right] \quad \left[\begin{array}{ccc} \text{Free Flow} & \text{Slowed} & \text{Start / Stop} \\ 0 & 0 & 0 \\ \beta_4 & 0 & 0 \\ 0 & \beta_5 & 0 \\ 0 & 0 & \beta_6 \\ \beta_4 & \beta_5 & 0 \\ \beta_4 & 0 & \beta_6 \\ 0 & \beta_5 & \beta_6 \\ \beta_4 & \beta_5 & \beta_6 \end{array} \right] \end{array} \right\}$$



Results for the 2^K model

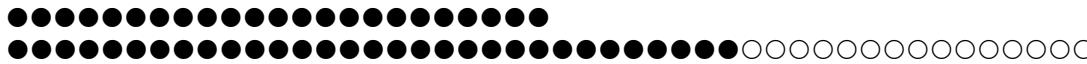




Latent Class Models with Cross Class Restrictions

$$\boldsymbol{\theta} = \begin{bmatrix} \text{Uncertainty} & \text{Toll Cost} & \text{Running Cost} \\ \beta_1 & \beta_2 & \beta_3 \end{bmatrix} \begin{bmatrix} \text{Free Flow} & \text{Slowed} & \text{Start / Stop} \\ 0 & 0 & 0 \\ \beta_4 & 0 & 0 \\ 0 & \beta_5 & 0 \\ 0 & 0 & \beta_6 \\ \beta_4 & \beta_5 & 0 \\ \beta_4 & 0 & \beta_6 \\ 0 & \beta_5 & \beta_6 \\ \beta_4 & \beta_5 & \beta_6 \end{bmatrix} \begin{bmatrix} \text{Prior Probs} \\ \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \\ \pi_7 \\ 1 - \sum_{j=1}^7 \pi_j \end{bmatrix}$$

- 8 Class Model: 6 structural utility parameters, 7 unrestricted prior probabilities.
- Reduced form has $8(6)+8 = 56$ parameters. ($\pi_j = \exp(a_j)/\sum_j \exp(a_j)$, $a_j = 0$.)
- EM Algorithm: Does not provide any means to impose cross class restrictions.
- “Bayesian” MCMC Methods: May be possible to force the restrictions – it will not be simple.
- Conventional Maximization: Simple



Practicalities

$$\beta = \begin{bmatrix} \gamma \\ c \end{bmatrix} = \begin{bmatrix} \text{Free parameters} \\ \text{Fixed values such as 0} \end{bmatrix}$$

$$\theta = K\beta = \begin{bmatrix} K_\gamma \\ K_c \end{bmatrix} \begin{bmatrix} \gamma \\ c \end{bmatrix}$$

Each row of K contains exactly one 1 and $M-1$ zeros

$$\log L = \sum_{i=1}^n \log L_i(\theta)$$

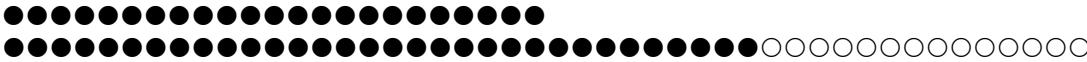
Maximization requires gradient and possibly Hessian wrt γ

These are simple sums of partials wrt θ .

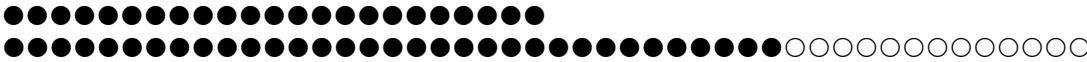
$$\frac{\partial \log L}{\partial \beta} = K' \frac{\partial \log L}{\partial \theta}, \quad \frac{\partial^2 \log L}{\partial \beta \partial \beta'} = K' \frac{\partial^2 \log L}{\partial \theta \partial \theta'} K$$

To complete the computation, discard rows wrt c .

Treat this as a conventional optimization problem.

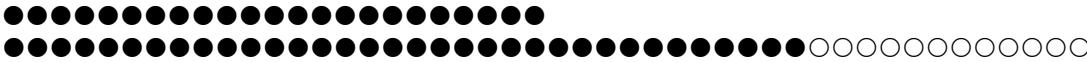


Latent Class Analysis of Nursing Home Cost Efficiency



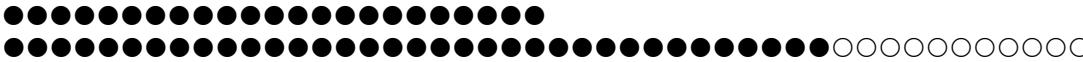
Latent Class Efficiency Studies

- Battese and Coelli – growing in weather “regimes” for Indonesian rice farmers
- Kumbhakar and Orea – cost structures for U.S. Banks
- Greene (Health Economics, 2005) – revisits WHO Year 2000 World Health Report



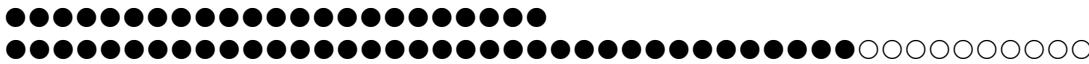
Studying Economic Efficiency in Health Care

- Hospital and Nursing Home
 - Cost efficiency
 - Role of quality (not studied today)
- AHRQ
- EWEPA, NAPW



Stochastic Frontier Analysis

- $\log C = f(\text{output, input prices, environment}) + v + u$
- $\varepsilon = v + u$
 - v = noise – the usual “disturbance”
 - u = inefficiency
- Frontier efficiency analysis
 - Estimate parameters of model
 - Estimate u (to the extent we are able – we use $E[u|\varepsilon]$)
 - Evaluate and compare observed firms in the sample



Unobserved quality of the management and efficiency measurement in the nursing home sector

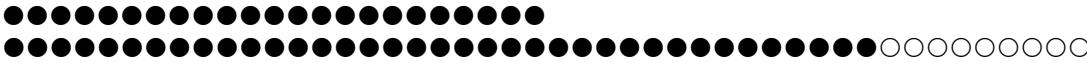
Massimo Filippini

and

William Greene

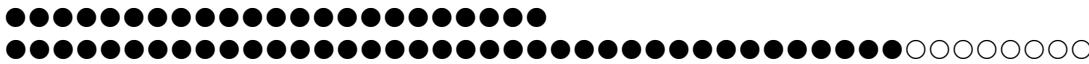
Departmeent of economics ETH
Zurich and Department of
Economics, University of Lugano,
Switzerland

Department of Economics, Stern
School of Business
New York University, USA

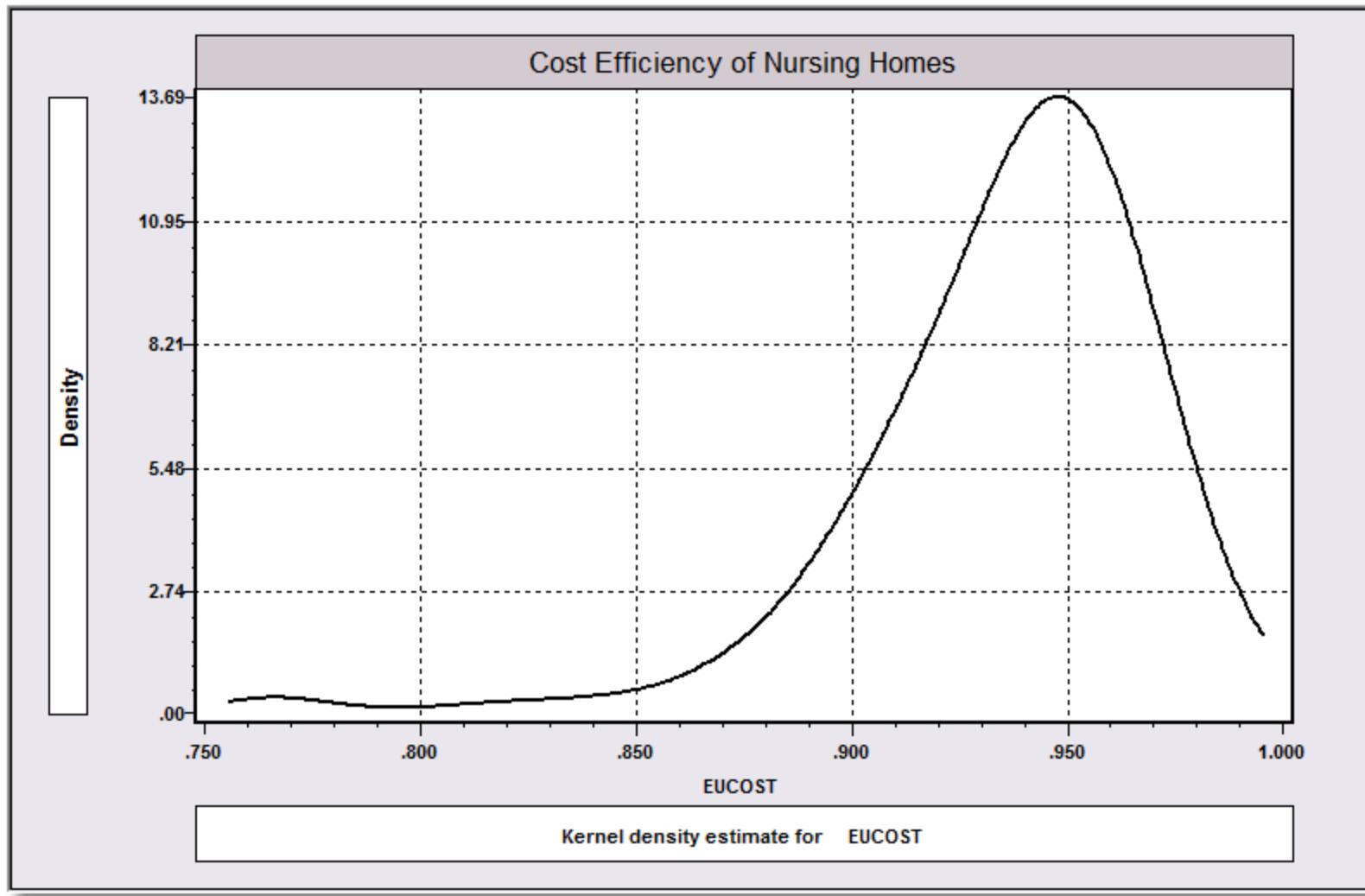


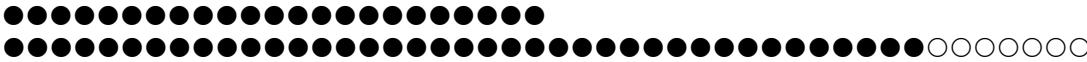
Nursing Home Costs

- 44 Swiss nursing homes, 13 years
- Cost, Pk, Pl, output, two environmental variables
- Estimate cost function
- Estimate inefficiency



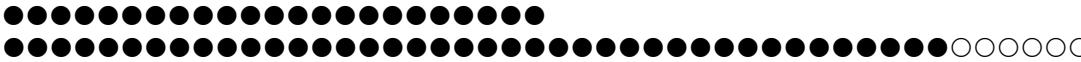
Estimated Cost Efficiency





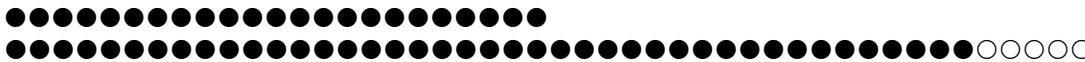
Inefficiency?

- Not all agree with the presence (or identifiability) of “inefficiency” in market outcomes data.
- Variation around the common production structure may all be nonsystematic and not controlled by management
- Implication, no inefficiency: $u = 0$.



A Two Class Model

- Class 1: With Inefficiency
 - $\log C = f(\text{output}, \text{input prices}, \text{environment}) + \sigma_v v + \sigma_u u$
- Class 2: Without Inefficiency
 - $\log C = f(\text{output}, \text{input prices}, \text{environment}) + \sigma_v v$
 - $\sigma_u = 0$
- Implement with a single zero restriction in a constrained (same cost function) two class model
- Parameterization: $\lambda = \sigma_u / \sigma_v = 0$ in class 2.



Latent Class / Panel Frontier Model
Dependent variable LNCT
Unbalanced panel has 44 individuals
Stochastic frontier (half normal model)
Model fit with 2 latent classes.

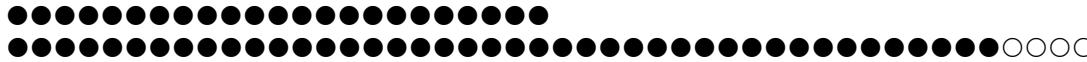
**LogL= 464 with a common frontier
model, 527 with two classes**

LNCT	Coefficient	Standard Error	z	Prob. z >Z*	95% Confidence Interval
Model parameters for latent class 1					
Constant	15.2268***	.00303	5030.79	.0000	15.2209 15.2328
LNQ1	.78749***	.03159	24.93	.0000	.72558 .84941
LNQ2	.29593***	.02387	12.40	.0000	.24914 .34272
LNPL	.76943***	.01443	53.33	.0000	.74115 .79771
LNY	.93837***	.00665	141.11	.0000	.92533 .95140
Sigma	.10758***	.00496	21.69	.0000	.09786 .11729
Lambda	1.90629***	.15870	12.01	.0000	1.59524 2.21735
Model parameters for latent class 2					
Constant	15.2268***	.00303	5030.79	.0000	15.2209 15.2328
LNQ1	.78749***	.03159	24.93	.0000	.72558 .84941
LNQ2	.29593***	.02387	12.40	.0000	.24914 .34272
LNPL	.76943***	.01443	53.33	.0000	.74115 .79771
LNY	.93837***	.00665	141.11	.0000	.92533 .95140
Sigma	.04997***	.00176	28.35	.0000	.04652 .05343
Lambda	0.0(Fixed Parameter).....			
Estimated prior probabilities for class membership					
Class1Pr	.29160***	.09619	3.03	.0024	.10307 .48014
Class2Pr	.70840***	.09619	7.36	.0000	.51986 .89693

Note: ***, **, * ==> Significance at 1%, 5%, 10% level.

Fixed parameter ... is constrained to equal the value or had a nonpositive st.error because of an earlier problem.

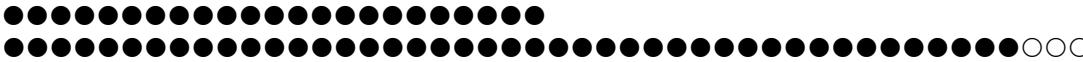
Class	Lambda	Sigma	Sigma(u)	Sigma(v)
1	1.906292	.107575	.095264	.049973
2	.000000	.049973	.000000	.049973



Revealing Additional Dimensions of Preference Heterogeneity in a Latent Class Mixed Multinomial Logit Model

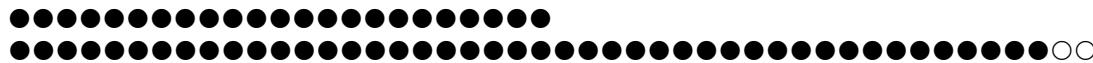
William H. Greene
Department of Economics
Stern School of Business
New York University, New York 10012
wgreene@stern.nyu.edu

David A. Hensher
Institute of Transport and Logistics Studies
Faculty of Economics and Business
The University of Sydney
NSW 2006 Australia
David.Hensher@sydney.edu.au



Freight Forwarding Experiment

- A stated choice (SC) framework within which a freight transporter defined a recent reference trip in terms of its time and cost attributes (detailed below), treating fuel as a separate cost item to the variable user charge (VUC),
- Whilst in-depth interviews and literature reviews revealed myriad attributes that influence freight decision making, we focussed on the subset of these attributes that were most likely to be directly affected by congestion charges.



Sydney Metropolitan Freight Stakeholders Study

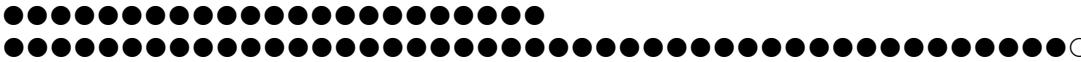


Practice Game

The alternatives on this screen represent three options for carrying out the freight trip you described - the trip as it occurred, and two trips involving new combinations of fuel taxes, distance-based congestion charges, and time and cost components. Please consider them and then answer the questions below:

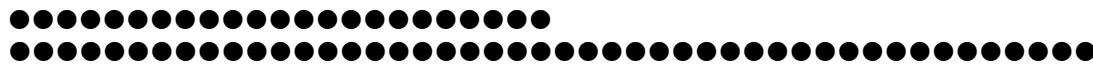
		Your Recent Trip	Trip Variation A	Trip Variation B
Free-flow travel time:	(definition)	15 minutes	19 minutes	22 minutes
Slowed-down travel time:	(definition)	55 minutes	28 minutes	82 minutes
Total time waiting to unload goods:		10 minutes	12 minutes	8 minutes
Likelihood of on-time arrival:		80%	70%	80%
Freight rate paid by the receiver of the goods:		\$450.00	\$461.67	\$461.67
Fuel cost:		\$15.57	\$19.46 (based on a 50% increase in fuel taxes)	\$23.35 (based on a 100% increase in fuel taxes)
Distance-based charges:		\$0.00	\$7.78	\$3.89
If your organisation and the receiver of the goods had to reach agreement on which alternative to choose, what would be your order of preference among alternatives? (please provide a choice for every alternative)	<p>My recent trip is</p> <p>My 1st choice</p> <p>My 2nd choice</p> <p>My 3rd choice</p> <p>Not acceptable</p>	<p>Trip Variation A is</p> <p>My 1st choice</p> <p>My 2nd choice</p> <p>My 3rd choice</p> <p>Not acceptable</p>	<p>Trip Variation B is</p> <p>My 1st choice</p> <p>My 2nd choice</p> <p>My 3rd choice</p> <p>Not acceptable</p>	
Which of these alternatives do you think would be acceptable to the receiver of the goods?			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Which alternative do you think the receiver of the goods would most prefer?		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Back](#)
[Trip Details](#)
[Relationship Details](#)
[Next](#)



LC-RP Model

- Latent Class Structure for Overall Model Format
- Random Parameters Model within Classes
- Prior class probabilities are heterogeneous



	Latent Class with decomposition of class membership probability			
	Fixed Parameters (M5)		Random Parameters (M6)	
Attributes:	Class 1	Class 2	Class 1	Class 2
Total time (mins)	0.0056 (0.69)	-0.00015 (-0.04)	-0.0069 (-1.96)	0.0015 (0.56)
On time delivery (percent)	-0.0062 (-0.21)	0.0279 (2.34)	-0.00056 ^{tp} (-0.040)	0.0276 ^{tp} (2.86)
Total cost (\$)	-0.0474 (-3.30)	-0.0043 (-1.91)	-0.0157 (-3.14)	-0.0055 (-4.26)
No variable charge dummy (1,0)	1.4220 (2.84)	-2.1567 (-1.34)	1.4109 (4.25)	-1.3472 (-2.95)
Class membership probability:	0.616	0.384	0.575	0.425
constant	1.1741 (3.30)	0	1.8200 (2.79)	0
Freight rate (\$)	-0.00076 (-2.32)	0	-0.0016 (-2.54)	0
BIC	1.7724		1.6498	
Log-likelihood	-372.85		-346.36	

Conclusion

Latent class modeling provides a rich, flexible platform for behavioral model building.

Thank you.